单个率Meta分析中率的不同转换方法应用现状

曾治宇

波科国际医疗贸易（上海）有限公司（北京 100020）

摘要 **目的** 考察当前单个率Meta分析的文献中对率的数据转换的实际使用情况。**方法** 在PubMed中检索2017年发表的单个率Meta分析的文献，从481条记录中筛选出145篇纳入分析。**结果** 在有全文的123篇文献中只有33篇（26.8%）文献交代了率的转换方法的使用，其中双重反正弦法20篇，logit转换8篇，平方根反正弦法3篇，对数转换1篇，直接使用原始率1篇。在这33篇文献中，率的转换方法的使用与汇总率的大小无关（P=0.217）。**结论** 单个率的Meta分析中率的转换方法是较为重要的因素，但各种转换方法的优劣尚无定论，发表的文献应加强对于率的数据转换等方法的说明。

关键词 单个率；Meta分析；数据转换

Current State of Data Transformation in Meta Analysis of Single Proportions

ZENG Zhi-yu

Boston Scientific Corporation BSC Int'l Medical Trading (Shanghai) Co., Ltd. Beijing, 100020, China

[Abstract] Objective To investigate the current status of data transformation in Meta analysis of single proportions. **Methods** A literature search in PubMed was performed to retrieve researches of Meta analysis of single proportions published in 2017, and 481 records were returned. **Results** In 123 researches with full texts, only 33 (26.8%) described data transformation for proportions. Among which, double arcsine transformation was used 20 times, logit

transformation 8, squared arcsine transformation 3, log transformation 1 and raw proportion 1. There was no relation between pooled portions and data transformation method (P=0.217). **Conclusion** Data transformation is important in Meta analysis of single proportions, however, it is yet to determined which transformation method is best. So, data transformation should be clarified in published papers.

Key words Single Proportions; Meta Analysis; Data Transformation

1 引言

Meta 分析是定量综合分析多个具有相同研究主题文献的过程，其本质是一种统计学方法[1]。Meta 分析的基本思想和方法产可追溯至 1904 年 Pearson 开始合并多个研究的结果，随后 Fisher 合并多个 P 值，但直到 1976 年 Glass GV 才首次命名了 Meta 分析（Meta-analysis）[2, 3]。

Meta 分析有多种类型，其中有一种特殊类型为单个率的 Meta 分析。单个率的 Meta 分析主要基于横断面研究的无对照二分类数据，这类数据常用于患病率、检出率、知晓率、病死率及感染率等的调查。这类数据的特点是仅有单组事件发生数和观察总数，而无对照组。本文关注的是这类数据的 Meta 分析，不包括以时间事件数据表达的发生率，也不包括诊断试验中的敏感性、特异性等效应量。

临床研究实践中对于率的数据处理，常用的方法是较为方便的正态近似法，但是当率或样本量偏小的情况下会出现较大的偏倚；并且，单个率的 Meta 分析采用倒方差法合并时，较小的率获得的权重可能会过大；此时应考虑率的不同的转换方法来满足统计分析的要求。目前单个率 Meta 分析的效应量合并在率的转换方法上有多种选择，如不转换（直接使用原始率）、对数转换、logit 转换、平方根反正弦转换（squired arcsine）及双重反正弦转换（double arcsine）等，但在实际操作中具体应在什么情况下转换，首选何种转换方法并不十分明确。本研究拟考察当前单个率 Meta 分析的文献中对率的数据转换的实际使用情况。

2 方法

2.1 文献检索

在 PubMed 中检索 2017 年发表的文献，使用的检索词包括：Meta analysis、prevalence、incidence、rate、percentage、frequency、proportion、mortality、complication 等。以自由词检索为主，注意检索词之间合理的逻辑组合。反复调整检索方案后，最终以((((meta[title] AND analysis[title]) AND (proportion[Title] OR prevalence[Title] OR percentage[Title] OR frequency[Title] OR incidence[Title] OR rate[Title] OR complication[title] OR mortality[title])) AND ("2017/01/01"[PDat] : "2017/12/31"[PDat]))) AND meta-analysis[Publication Type]获得 481 条记录（2018 年 6 月 17 日检索）。

2.2 文献筛选

将 481 条记录导入 NoteExpress 文献管理软件，全部阅读标题及摘要，筛选标准为摘要中出现以百分比表达的 Meta 分析汇总的率，进而获取全文。文献筛选流程见图 2，最终纳入 145 篇，其中 123 获得全文。

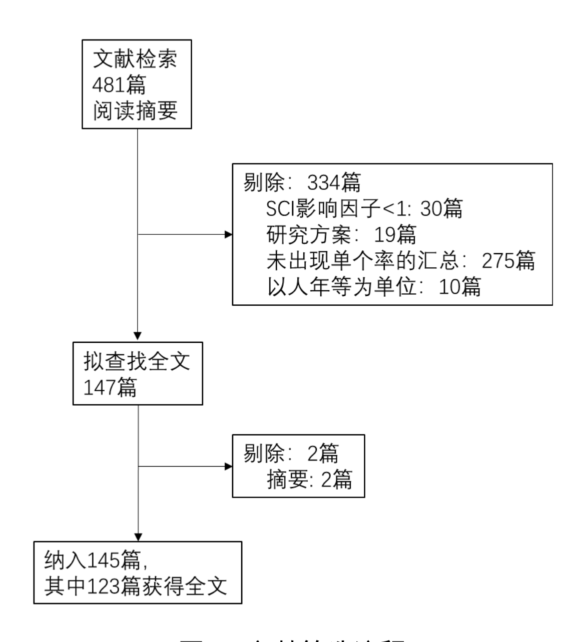


图 1 文献筛选流程

2.3 数据提取

提取率的类型、研究个数、最小样本量、Meta 分析软件、率的转换方法、汇总率及 I^2 值等参数，详见 Error: Reference source not found。当一篇文献有多次单个率的 Meta 分析时，选择最主要的一个，或者最先出现的那个。

2.4 统计分析

主要是统计描述，计数资料以数值及百分比表达，计量资料以均数±标准差表达，同时给出中位数。无正式的假设检验，对汇总率的大小与率的转换方法作简单线性回归分析。统计软件为R（版本3.5.2）。

3 结果

在入选的145篇文献中，来自中国大陆的文献最多，37篇（25.5%），接下来是美国17篇（11.7%），加拿大15篇（10.3%），英国13篇（9.0%）。绝大多数文献使用了prevalence一词，122篇（84.1%），其他还有incidence 11篇（7.6%），rate 5篇（3.4%），proportion 3篇（2.1%）等。需要注意的是，使用prevalence词语的文献中有些是发生率的概念，但并没有用类似“人年数”的单位作分母；而有些文献虽然使用了incidence，但其实是prevalence，因统计方法相同，故都纳入了分析。

每个Meta分析纳入的研究个数从2-184不等，平均38.2±26.2，中位数28（16-49）。最小样本量从2-39211不等，平均628.3±4202.3，中位数49（20-134）。汇总的率从0.0003-0.984不等，平均0.245±0.227，中位数0.179（0.070-0.351）。

在有全文的123篇文献中，Meta分析的软件以STATA使用最多，49篇（39.8%），Comprehensive Meta-Analysis（CMA）26篇（21.1%），R 21篇（17.1%），其他还有MetaXL 4篇（3.3%），Open Meta-Analysis 3篇（2.4%），MedCalc 3篇（2.4%），StatsDirect 3篇（2.4%）等。

在有全文的123篇文献中只有33篇（26.8%）文献交代了率的转换方法的使用，双重反正弦法20篇，logit转换8篇，平方根反正弦法3篇，对数转换1篇，直接使用原始率1篇。只有1篇文献（编号111）[4]比较了5种数据转换后的正态性，最后选择了平方根反正弦法。在这33篇文献中，率的转换方法的使用与汇总率的大小无关（简单线性回归，P=0.217）。

只有5篇文献交代了率的汇总的权重，其中4篇为倒方差，1篇为样本量。仅2篇文献提到当事件数为0时增加0.5进行校正。绝大多数文献给出了I^2值，且反应了单个率Meta分析中各研究间明显的异质性，也多采用了随机效应模型。

4 讨论

本文对单个率的 Meta 分析中率的转换方法的应用现状作了较为细致的描述，使用最多的方法为双重反正弦法，其次为 logit 转换和平方根反正弦法。但总体看来，发表的文献对于率的转换的交代不够充分，对于权重、校正等说明也明显欠缺。

在单个率 Meta 分析的背景下，初步检索国内文献有一些单个率 Meta 分析软件使用的介绍[5-8]，如王佩鑫等介绍了无对照二分类资料的 Meta 分析方法及 Stata 实现[5]，罗美玲等介绍了在 R 软件中实现单个率的 Meta 分析[6]，肖静等介绍了单一率 Meta 分析的 Matlab 软件实现[7]，陈月红等介绍了无对照二分类数据的 Meta 分析在 RevMan 软件中的实现[8]。但这些文献基本为介绍性，重点在与各种软件的使用方法，并没有对率的不同转换方法进行比较。

检索国外文献，类似研究也不多见[9-11]。Trikalinos 等[9]对单个比例和发生率的 Meta 分析分别进行了进行模拟研究，在单个比例的 Meta 分析中，根据比例的分布、均值、方差、研究个数及研究样本量等构建了 700 余种 Meta 分析的情形，对每种情形进行 1000 次随机抽样模拟，同时考察固定效应模型和随机效应模型。 率的转化包括了不转换、logit 转换和反正弦转换等 3 种方法，结果显示反正弦转换是较好的方法。Barendregt 等[10]对患病率（Prevalence）的 Meta 分析进行了模拟研究，假设了样本量为 20-180 的 9 个研究，设定患病率为 0.05，随机效应模型时假设患病率在不同的研究中为正态分布，均值为 0.05，标准差为 0.005。进行 1000 次随机抽样，率的转换方法包括了不转换、logit 转换及双重反正弦转换等 3 种方法，结果双重反正弦法表现最佳。遗憾的是这两个研究各自只比较了 3 种率的转换方法，且没有反正弦转换与双重反正弦转换的直接对比。而令人担忧的是最近一篇文献认为在单个率的 Meta 分析中使用双重反正弦转换有严重缺陷，并建议使用广义线性混合模型[11]。

Meta 分析的方法学研究具有重要意义，特别是在率较小的时候。例如 2007 年新英格兰医学杂志发表了一篇糖尿病药物罗格列酮安全性的 Meta 分析[12]，认为罗格列酮增加心肌梗死的风险，并且有增加心血管死亡的趋势。这篇文章一经发表就引起了广泛的争议，在 Meta 分析的方法学上也有不同的看法，特别是针对较小的率的处理方法，不同的 Meta 分析甚至得到相反的结果[13-17]。尽管如

此，新英格兰医学杂志发表的这篇 Meta 分析成为药物安全性问题的重要依据，促进了欧洲于 2010 年将该药下市，美国 FDA 于 2011 年严格限制了该药的使用。到后来有了更多的临床依据之后，美国 FDA 于 2013 年撤销了药物使用的限制，2015 年再次重申了药物安全的观点[18]。

当前对于有对照组的小概率事件的 Meta 分析已有较多的探讨[19-23]，但对于单个率 Meta 分析率的不同转换方法适用性的研究仍然不够充分[9, 10]。临床上对单个率进行汇总分析需求越来越多，包括药物或器械的不良事件分析等。在这样的一个背景下，Meta 分析应当对方法学有较为详细的介绍，保证 Meta 分析的可重复性。

本文的局限在于：1）文献检索尚不够充分全面，仅检索了 PubMed 一年的文献，尚有部分文献未获得全文。但研究结果应当已经能反应当前单个率 Meta 分析方法学交代尚不充分的基本事实。2）当一篇文献有多个单个率的 Meta 分析时，只选择了最主要的或最早出现的分析。3）重点关注了率的转换方法的应用，对于其他一些重要的参数如权重、校正等仅略有提及，特别是 95%可信区间的构造方法未能包含，但现有的信息也反映了这些参数在发表的文献中很少提及。

总之，单个率的 Meta 分析中率的转换方法是较为重要的因素，但各种方法的优劣尚无定论，除进一步加强方法学的研究外，发表的文献也应注重对于率的转换等方法的说明。

5 参考文献

[1] 李幼平. 循证医学[M]. 北京:人民卫生出版社,2014.

[2] Gurevitch J, Koricheva J, Nakagawa S, et al. Meta-analysis and the science of research synthesis[J]. Nature,2018,555(7695):175-182.

[3] Glass G V. Primary, secondary, and meta-analysis of research[J]. Educ. Res.,1976,5(10):3-8.

[4] Wang T, Liu Y, Li Z, et al. Prevalence of intimate partner violence (IPV) during pregnancy in China: A systematic review and meta-analysis[J]. PLoS One,2017,12(10):e175108.

[5] 王佩鑫，李宏田，刘建蒙. 无对照二分类资料的 Meta 分析方法及 Stata 实现[J]. 循证医学,2012,12(01):52-55.

[6] 罗美玲，谭红专，周权，等. 在 R 软件中实现单个率的 Meta 分析[J]. 循证医学,2013,13(03):181-184.

[7] 肖静，吴桂云，高月霞，等. 单一率 Meta 分析的 Matlab 软件实现[J]. 中国卫生统计,2014,31(04):715-717.

[8] 陈月红，杜亮，耿兴远，等. 无对照二分类数据的 Meta 分析在 RevMan 软件中的实现[J]. 中国循证医学杂志,2014,14(07):889-896.

[9] Trikalinos T A, Trow P, Schmid C H. Simulation-Based Comparison of Methods for Meta-

Analysis of Proportions and Rates[M]. Rockville (MD):Agency for Healthcare Research and Quality (US),2013.
[10] Barendregt J J, Doi S A, Lee Y Y, et al. Meta-analysis of prevalence[J]. Journal of Epidemiology and Community Health,2013,67(11):974-978.
[11] Schwarzer G, Chemaitelly H, Abu-Raddad L J, et al. Seriously misleading results using inverse of freeman-tukey double arcsine transformation in meta-analysis of single proportions[J]. Res Synth Methods,2019 doi: 10.1002/jrsm.1348. [Epub ahead of print].
[12] Nissen S E, Wolski K. Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes[J]. New England Journal of Medicine,2007,356(24):2457-2471.
[13] 卢宇，马德琳，余学锋. 罗格列酮和吡格列酮的心血管安全性[J]. 药物不良反应杂志,2015(2).
[14] 陈燕铭，唐喜香，曾龙驿. 再评罗格列酮的心血管安全性[J]. 中华内分泌代谢杂志,2014,30(6).
[15] Dahabreh I J. Meta-analysis of rare events: an update and sensitivity analysis of cardiovascular events in randomized trials of rosiglitazone[J]. Clinical Trials: Journal of the Society for Clinical Trials,2008,5(2):116-120.
[16] Cai T, Parast L, Ryan L. Meta-analysis for rare events[J]. Statistics in Medicine,2010,29(20):2078-2089.
[17] Nissen S E, Wolski K. Rosiglitazone Revisited[J]. Archives of Internal Medicine,2010,170(14):1191-1201.
[18] FDA Drug Safety Communication: FDA eliminates the Risk Evaluation and Mitigation Strategy (REMS) for rosiglitazone-containing diabetes medicines[Z]. 2015:2018.
[19] Bhaumik D K, Amatya A, Normand S T, et al. Meta-Analysis of Rare Binary Adverse Event Data[J]. Journal of the American Statistical Association,2012,107(498):555-567.
[20] Ma Y, Chu H, Mazumdar M. Meta-analysis of Proportions of Rare Events－A Comparison of Exact Likelihood Methods with Robust Variance Estimation[J]. Communications in Statistics - Simulation and Computation,2015,45(8):3036-3052.
[21] Bakbergenuly I, Kulinskaya E, Morgenthaler S. Inference for binomial probability based on dependent Bernoulli random variables with applications to meta-analysis and group level studies[J]. Biometrical Journal,2016,58(4):896-914.
[22] Bai O, Chen M, Wang X. Bayesian Estimation and Testing in Random Effects Meta-Analysis of Rare Binary Adverse Events[J]. Statistics in Biopharmaceutical Research,2016,8(1):49-59.
[23] Jackson D, Law M, Stijnen T, et al. A comparison of seven random-effects models for meta-analyses that estimate the summary odds ratio[J]. Stat Med,2018,37(7):1059-1085.